

# POINTWISE EQUIDISTRIBUTION AND TRANSLATES OF MEASURES ON HOMOGENEOUS SPACES

OSAMA KHALIL

**ABSTRACT.** Let  $(X, \mathfrak{B}, \mu)$  be a Borel probability space. Let  $T_n : X \rightarrow X$  be a sequence of continuous transformations on  $X$ . Let  $\nu$  be a probability measure on  $X$  such that  $\frac{1}{N} \sum_{n=1}^N (T_n)_* \nu \rightarrow \mu$  in the weak-\* topology. Under general conditions, we show that for  $\nu$  almost every  $x \in X$ , the measures  $\frac{1}{N} \sum_{n=1}^N \delta_{T_n x}$  get equidistributed towards  $\mu$  if  $N$  is restricted to a set of full upper density. We present applications of these results to expanding translates of curves on homogeneous spaces and translates of orbits of symmetric groups. As a corollary, we prove equidistribution of certain sparse orbits of the horocycle flow on quotients of  $SL(2, \mathbb{R})$ , starting from every point in almost every direction.

## 1. INTRODUCTION

Many problems in number theory and geometry can be recast in terms of the equidistribution of translates of appropriate measures on quotients of certain Lie groups. The general set up of these results is a Borel probability space  $(X, \mathfrak{B}, \mu)$ , a probability measure  $\nu$  on  $X$  (usually singular with respect to  $\mu$ ) and a sequence of transformations  $T_n : X \rightarrow X$  such that

$$\frac{1}{N} \sum_{n=1}^N (T_n)_* \nu \xrightarrow{N \rightarrow \infty} \mu \quad (1.1)$$

where  $(T_n)_* \nu$  is the pushforward of  $\nu$  under  $T_n$  and the convergence is in the weak-\* topology. A natural question is to what extent can one extend such results to describe the behavior of measures of the form

$$\frac{1}{N} \sum_{n=1}^N \delta_{T_n x} \quad (1.2)$$

where  $\delta_y$  denotes the dirac delta measure at a point  $y$ .

Recently<sup>1</sup>, this question was addressed by Chaika and Eskin [CE] in the context of flat surfaces. There,  $X$  is some affine submanifold of the moduli space of flat structures on a surface,  $\mu$  is a natural affine  $SL(2, \mathbb{R})$  invariant measure,  $\nu$  is the measure supported on an orbit of  $SO(2)$  and  $T_n = a(n) = \text{diag}(e^n, e^{-n})$ . They show that for  $\nu$  almost every  $x$ , the measures in (1.2) get equidistributed towards  $\mu$ .

In the context of homogeneous spaces, Shi [Shi] explored this question for translates of measures supported on (pieces of) orbits of certain horospherical subgroups of Lie groups by one parameter diagonalizable subgroups. Here  $X$  is a homogeneous space for a Lie group  $G$ ,  $\nu$  is a measure on an orbit of a certain horospherical subgroup,  $T$  is an Ad-diagonalizable

<sup>1</sup>The following results were stated for flows, but we mention their equivalent discrete analogues for consistency with our set up.

element of  $G$  and  $T_n = T^n$ . Equidistribution of measures of the form (1.2) towards the natural  $G$ -invariant Haar measure  $\nu$  almost everywhere is proven.

In [KSW], an effective version of this result is obtained via different methods. The convergence of measures of the form (1.2) is proven for general dynamical systems under the hypothesis of some form of exponential mixing of the transformation  $T$  with respect to the non-invariant measure  $\nu$ .

In all three cases, equidistribution was obtained by exploiting specific properties of the system at hand, while not directly utilizing the fact that (1.1) holds.

**1.1. Statement of Results.** In this article, we approach this question in the general context of continuous measure preserving transformations, assuming (1.1) only. We obtain equidistribution results of measures in (1.2) under general conditions yet only along subsequences of full upper density. Recall the definition of upper density:

**Definition.** The upper density of a subset  $A \subseteq \mathbb{N}$ , denoted by  $\overline{d}(A)$  is defined to be

$$\overline{d}(A) = \limsup_{N \rightarrow \infty} \frac{\#(A \cap [1, N])}{N}$$

In what follows,  $X$  will be a locally compact, second countable topological space and  $\mathcal{B}$  is its Borel  $\sigma$ -algebra. A pair  $(X, \mathcal{B})$  will be called a standard Borel space. The following is our first main result for the case when translations are done by powers of a single transformation.

**Theorem 1.1.** *Let  $(X, \mathcal{B}, \mu)$  be a standard Borel probability space. Let  $T$  be an ergodic continuous measure preserving transformation on  $X$ . Let  $\nu$  be a probability measure on  $X$  such that*

$$\frac{1}{N} \sum_{n=0}^{N-1} T_*^n \nu \xrightarrow[N \rightarrow \infty]{\text{weak-}^*} \mu$$

*Then, for  $\nu$ -almost every  $x \in X$ , there exists a sequence  $A(x) \subseteq \mathbb{N}$ , of upper density 1, such that for all  $\psi \in C_c(X)$ ,*

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(x)}} \frac{1}{N} \sum_{n=0}^{N-1} \psi(T^n x) = \int \psi d\mu$$

The proof of Theorem 1.1 relies on an adaptation of the weak-type maximal inequality and follows similar lines to the proof of the classical Birkhoff ergodic theorem.

Our next result concerns the more general situation of translating by sequences of transformations. Unfortunately, in this generality we are only able to obtain equidistribution under more conditions on the measure  $\mu$  which are meant to abstract the necessary features in the set up of homogeneous spaces and flat surfaces driving equidistribution. It would be interesting to know if a similar statement is still valid under weaker, more general hypotheses.

**Theorem 1.2.** *Let  $(X, \mathcal{B}, \mu)$  be a standard Borel probability space. Let  $(T_n)_n$  be a sequence of continuous transformations on  $X$ . Let  $S : X \rightarrow X$  be an ergodic continuous  $\mu$  preserving transformation. Let  $\nu$  be a probability measure on  $X$ . Assume the following holds:*

- (1)  $\frac{1}{N} \sum_{n=1}^N (T_n)_* \nu \xrightarrow[N \rightarrow \infty]{} \mu$  in the weak-\* topology.
- (2) For  $\nu$ -almost every  $x \in X$ , any limit point of the sequence  $\frac{1}{N} \sum_{n=1}^N \delta_{T_n x}$  is  $S$ -invariant, where  $\delta_y$  denotes the dirac delta measure at  $y$ .
- (3) There exists a Borel measurable set  $Z \in \mathcal{B}$  such that

- (a)  $\mu(Z) = 0$  and for all ergodic  $S$  invariant measures  $\lambda \neq \mu$ ,  $\lambda(Z) = 1$ .
- (b)  $Z = \cup_n K_n$ , where  $K_n \subseteq K_{n+1}$  and  $K_n$  is a compact set for all  $n$ .

Then, for  $\nu$ -almost every  $x \in X$ , there exists a sequence  $A(x) \subseteq \mathbb{N}$ , of upper density equal to 1, such that for all  $\psi \in C_c(X)$ ,

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(x)}} \frac{1}{N} \sum_{n=1}^N \psi(T_n x) = \int \psi d\mu$$

We now discuss some applications of these results. We will need to set up some notation first.

**1.2. Expanding Translates of Curves.** In a series of articles, Shah proved that the translates of certain non-degenerate curves on homogeneous spaces, by appropriate diagonal elements, get equidistributed. These results for certain curves on quotients of  $SL(n, \mathbb{R})$  were utilized to derive almost everywhere Dirichlet non-improvability statements with respect to the Lebesgue measure on such curves. See [Sha1, Sha2, Sha3, SY] for a more thorough treatment.

Theorem 1.1 applies to this situation as well. For concreteness, we'll limit ourselves to the situation considered in [SY]. The other situations follow similarly.

Let  $m, n \in \mathbb{N}$ . Let  $G = SL(m+n, \mathbb{R})$ . Let  $M(m \times n, \mathbb{R})$  be the space of  $m$  by  $n$  real matrices. Let

$$\varphi : [0, 1] \rightarrow M(m \times n, \mathbb{R})$$

be an analytic curve. For  $X \in M(m \times n, \mathbb{R})$  and  $t \in \mathbb{R}$ , define

$$u(X) = \begin{pmatrix} I_m & X \\ 0 & I_n \end{pmatrix}, \quad a(t) = \begin{pmatrix} e^{nt} I_m & 0 \\ 0 & e^{-mt} I_n \end{pmatrix}$$

where  $I_m$  and  $I_n$  denote the identity matrices in dimensions  $m$  and  $n$  respectively. Then,  $u(X) \in G$  and  $\varphi$  can be regarded as a curve in  $G$ .

Shah introduced a notion of generic curves which is an appropriate form of non-degeneracy in this context. This notion was refined in [SY] to the notion of supergeneric curves. For our purposes, we will assume that  $\varphi$  is a supergeneric curve. See [SY], Definition 1.1.

Let  $L$  be a Lie group containing  $G$  and let  $\Gamma$  be a lattice in  $L$ . Let  $x \in L/\Gamma$  be such that  $\overline{Gx} = L/\Gamma$ . Let  $\mu_L$  be the unique  $L$  invariant Haar probability measure on  $L/\Gamma$ . Let  $\nu$  be the pushforward of the Lebesgue measure on  $[0, 1]$  under  $s \mapsto u(\varphi(s))x \in L/\Gamma$  as defined above.

The following theorem is an application of Theorem 1.1 in this set up.

**Theorem 1.3.** *In the notation above, if  $\varphi$  is a supergeneric curve, then for  $\nu$  almost every  $y = u(\varphi(s))x \in L/\Gamma$ , there exists a set  $A(s) \subseteq \mathbb{N}$  of upper density 1 such that*

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(s)}} \frac{1}{N} \sum_{n=1}^N \delta_{a(n)u(\varphi(s))x} = \mu_L$$

*Proof.* By Theorem 1.4 in [SY],  $a(n)_* \nu \rightarrow \mu_L$  as  $N \rightarrow \infty$ . Hence, the conclusion follows by Theorem 1.1  $\square$

**1.3. Translates of Orbits of Symmetric Groups.** Our next application involves applying Theorem 1.2 to translates of closed orbits of symmetric groups. As a consequence of Theorem 1.4 below, we obtain an equidistribution result for sparse orbits of the horocycle flow. We will need some notation to state our results precisely.

Let  $G$  be a connected semisimple Lie group with finite center and let  $\Gamma$  be a lattice in  $G$ . Let  $\mu_{G/\Gamma}$  denote the unique  $G$  invariant Haar probability measure on  $G/\Gamma$ . Let  $H$  be a closed subgroup such that  $G/H$  is an affine symmetric space i.e.  $H$  is the fixed point set of an involution of  $G$ .

Assume that  $H \cap \Gamma$  is a lattice in  $H$  and let  $\mu_H$  denote the unique left  $H$ -invariant Haar measure supported on the closed orbit  $H\Gamma/\Gamma \subset G/\Gamma$ , normalized to be a probability measure.

Our main theorem in this set up is the following. We state it imprecisely here and we refer the reader to Theorem 6.1 below for the precise statement.

**Theorem 1.4.** *Let  $g_n$  be a sequence tending to infinity in  $G/H$  and satisfying certain growth conditions. Then, after possibly passing to a subsequence of  $g_n$ , also denoted by  $g_n$ , for  $\mu_H$  almost every  $x \in G/\Gamma$ , there exists a sequence  $A(x) \subseteq \mathbb{N}$  of upper density 1 such that*

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(x)}} \frac{1}{N} \sum_{n=1}^N \delta_{g_n x} = \mu_{G/\Gamma}$$

Apart from Theorem 1.2, a key ingredient in the proof of Theorem 1.4 is an adaptation of a lemma due to Chaika and Eskin [CE] to this set up. The growth conditions referred to in the statement are needed for this step. This lemma shows that for  $\nu$  almost every  $x$ , any limit point of the measures considered in Theorem 1.4 is invariant under a one parameter unipotent subgroup. Passage to a subsequence of  $g_n$  is needed here to insure that limiting measures are invariant by a limit one-parameter unipotent subgroup. This allows us to use Ratner's measure classification theorem to verify the hypotheses of Theorem 1.2.

**1.3.1. Sparse Equidistribution.** A resistant question in homogeneous dynamics is the one concerning sparse equidistribution of the horocycle flow. Recent progress was achieved in [Ven] for sequences of the form  $n^{1+\gamma}$  for small values of  $\gamma$ . See also [TV, Zhe, FFT] for more results in this direction and the work of Sarnak and Ubis on the equidistribution along the primes [SU]. In [Kat], the Hausdorff dimension of the "exceptional set" was shown to be not full. An interesting consequence of Theorem 1.4 and its proof is the following modest contribution to this question.

**Corollary 1.5.** *Let  $G = SL(2, \mathbb{R})$ ,  $\Gamma \subset G$  a lattice and let  $K = SO(2)$ . Let  $\lambda > 0$  and for  $n \in \mathbb{N}$ , let  $t_n = e^{\lambda n}$ . Let*

$$g_n = \begin{pmatrix} 1 & t_n \\ 0 & 1 \end{pmatrix}, k_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

*Then, for every  $x \in G/\Gamma$  and for almost every  $\theta \in [0, 2\pi]$ , there exists a sequence  $A(\theta) \subseteq \mathbb{N}$  of upper density 1 such that*

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(\theta)}} \frac{1}{N} \sum_{n=1}^N \delta_{g_n k_\theta x} = \mu_{G/\Gamma}$$

*Moreover, if  $G/\Gamma$  is compact then  $A(\theta) = \mathbb{N}$ .*

The main point of this corollary is that it holds for every  $x$ . This is not guaranteed by any general theorem on sparse equidistribution almost everywhere. The growth conditions required by Theorem 1.4 limit us to horocycle orbits along (sub)exponentially sparse sequences (see remark after Theorem 6.1). In particular, any improvement on the bounds in Lemma 5.8 below will automatically yield results for less sparse sequences.

The paper is organized as follows. In § 2, we prove an analogue of the maximal ergodic theorem in our set up. We use this to prove Theorems 1.1 and 1.2 in § 3 and § 4. In § 5, we show invariance of the measures considered in Theorem 1.4 by one-parameter unipotent subgroups. In § 6, we provide a proof of Theorem 1.4 and Corollary 1.5.

**Acknowledgement.** I would like to thank my advisor, Nimish Shah, for suggesting the problem and his guidance, support and for numerous valuable discussions.

## 2. AN ANALOGUE OF THE MAXIMAL INEQUALITY

The following proposition is an extension of the classical maximal ergodic theorem to the set up involving sequences of transformations and a non-invariant measure.

**Proposition 2.1.** *Let  $(X, \mathcal{B})$  be a standard Borel probability space. Let  $T_n$  be a sequence of continuous transformations on  $X$ . Let  $\nu, \mu$  be probability measures on  $X$  such that*

$$\frac{1}{N} \sum_{n=0}^{N-1} (T_n)_* \nu \xrightarrow[N \rightarrow \infty]{\text{weak-}^*} \mu$$

*Let  $\psi \in C_c(X)$  and let  $\alpha > 0$  and  $\beta \in (0, 1)$ . For every  $j, N \geq 1$ , define the set*

$$E_{\alpha, N, j}^{\psi} = \left\{ x \in X : \sup_{1 \leq M \leq N} \left| \frac{1}{M} \sum_{k=j}^{j+M-1} \psi(T_k x) \right| > \alpha \right\}$$

*Then, for every  $N \gg 1$ , depending on  $\psi$ , there exists some  $0 \leq j_N < \beta N$ , such that*

$$\alpha \beta \nu(E_{\alpha, N, j_N}^{\psi}) \leq 12 \|\psi\|_{L^1(\mu)}$$

We will deduce this proposition from the classical maximal inequality for  $l^1(\mathbb{Z})$  which is a consequence of Vitali's covering lemma. The proof follows closely the proof of the usual maximal inequality.

**Lemma 2.2** (Lemma 2.29, [EW]). *Let  $\phi \in l^1(\mathbb{Z})$ . Define the following maximal function, for  $a \in \mathbb{Z}$ :*

$$\phi^*(a) = \sup_{N \geq 1} \left| \frac{1}{N} \sum_{i=0}^{N-1} \phi(i+a) \right|$$

*Let  $\alpha > 0$  and define*

$$E_{\alpha} = \{a \in \mathbb{Z} \mid \phi^*(a) > \alpha\}$$

*Then,*

$$\alpha |E_{\alpha}| \leq 3 \|\phi\|_{l^1(\mathbb{Z})}$$

*Proof of Proposition 2.1.* Let  $\psi \in C_c(X)$  and let  $\alpha > 0$ ,  $\beta \in (0, 1)$ . Let  $N \geq 1$  and let  $E_{\alpha, N, j}$  be as in the statement. Let  $x \in X$  and let  $J > N$  be a parameter to be determined later. Define the following function

$$\phi(j) = \begin{cases} \psi(T_j x) & 0 \leq j \leq J \\ 0 & \text{otherwise} \end{cases}$$

Then, clearly  $\phi \in l^1(\mathbb{Z})$ . For  $a \in \mathbb{Z}$ , define the following two functions

$$\phi^*(a) = \sup_{1 \leq M} \left| \frac{1}{M} \sum_{k=0}^{M-1} \phi(k+a) \right|, \quad \phi_N^*(a) = \sup_{1 \leq n \leq N} \left| \frac{1}{n} \sum_{k=0}^{n-1} \phi(k+a) \right|$$

Define the corresponding exceptional sets

$$E_\alpha^\phi = \{a \in \mathbb{Z} \mid \phi^*(a) > \alpha\}, \quad E_{\alpha, N}^\phi = \{a \in [0, J - N - 1] \mid \phi_N^*(a) > \alpha\}$$

By Lemma 2.2 applied to  $\phi$ , we have

$$\alpha |E_{\alpha, N}^\phi| \leq \alpha |E_\alpha^\phi| \leq 3 \|\phi\|_{l^1(\mathbb{Z})} \quad (2.1)$$

Note that for  $a \in [0, J - N - 1]$ , we have

$$\phi_N^*(a) = \sup_{1 \leq n \leq N} \left| \frac{1}{n} \sum_{k=0}^{n-1} \psi(T_{k+a} x) \right| \quad (2.2)$$

Let  $\chi_{\alpha, N, j}$  denote the indicator function of  $E_{\alpha, N, j}^\psi$ . Thus, for  $j \in [0, J - N - 1]$ ,

$$\chi_{\alpha, N, j}(x) = 1 \text{ if and only if } j \in E_{\alpha, N}^\phi \quad (2.3)$$

Thus, combining (2.1), (2.2) and (2.3) along with the definition of  $\phi$ , we get

$$\alpha \sum_{j=0}^{J-N-1} \chi_{\alpha, N, j}(x) = \alpha |E_{\alpha, N}^\phi| \leq 3 \sum_{j=0}^J |\phi(j)| = 3 \sum_{j=0}^J |\psi(T_j x)|$$

Integrating both sides of the above with respect to  $\nu$  yields

$$\alpha \sum_{j=0}^{J-N-1} \nu(E_{\alpha, N, j}^\psi) \leq 3 \sum_{j=0}^J \int |\psi(T_j x)| d\nu(x) \quad (2.4)$$

Taking  $J = (1 + \beta)N$  in (2.4) and dividing both sides by  $J - N$

$$\begin{aligned} \frac{\alpha}{\beta N} \sum_{j=0}^{\beta N-1} \nu(E_{\alpha, N, j}^\psi) &\leq 3 \frac{(1 + \beta)N + 1}{\beta N} \frac{1}{(1 + \beta)N + 1} \sum_{j=0}^{(1+\beta)N} \int |\psi(T_j x)| d\nu(x) \\ &\leq \frac{6}{\beta} \frac{1}{(1 + \beta)N + 1} \sum_{j=0}^{(1+\beta)N} \int |\psi(T_j x)| d\nu(x) \end{aligned} \quad (2.5)$$

Now, by assumption,

$$\frac{1}{(1 + \beta)N + 1} \sum_{j=0}^{(1+\beta)N} \int |\psi(T_j x)| d\nu(x) \rightarrow \int |\psi| d\mu = \|\psi\|_{L^1(\mu)}$$

Thus, for all  $N$  sufficiently large, depending only on  $\psi$ , we have

$$\frac{1}{(1+\beta)N+1} \sum_{j=0}^{(1+\beta)N} \int |\psi(T_j x)| d\nu(x) \leq 2\|\psi\|_{L^1(\mu)}$$

Combining this with (2.5), we get for all  $N$  sufficiently large,

$$\frac{1}{\beta N} \sum_{j=0}^{\beta N-1} \nu(E_{\alpha, N, j}^\psi) \leq \frac{12\|\psi\|_{L^1(\mu)}}{\alpha\beta}$$

Thus, there must exist some  $j = j(N) \in [0, \beta N - 1]$  for which the conclusion of the Proposition holds. □

### 3. AN ANALOGUE OF BIRKHOFF'S ERGODIC THEOREM - THEOREM 1.1

This section is dedicated to the proof of Theorem 1.1. With the maximal inequality for the non-invariant measure  $\nu$  in place (Proposition 2.1), our proof will follow the same lines as Bourgain's approach to deduce the classical Birkhoff theorem from the mean ergodic theorem (cf. [Bou], Section 2-C).

Recall that for a sequence of sets  $A_n$ , the *limsup* of these sets is the set of elements which belong to  $A_n$  for infinitely many  $n$ . More precisely,

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{n \geq 1} \bigcup_{k \geq n} A_k$$

The following simple observation will be used repeatedly in what follows.

**Lemma 3.1.** *Let  $X$  be a standard Borel space and let  $\mu$  be a probability measure on  $X$ . Let  $A_n \subseteq X$  be a sequence of measurable sets such that  $\mu(A_n) \geq \alpha$  for some  $\alpha \in [0, 1]$ . Then,*

$$\mu\left(\limsup_{n \rightarrow \infty} A_n\right) \geq \alpha$$

*Proof.* This follows from the definition of  $\limsup_{n \rightarrow \infty} A_n$  as a decreasing intersection and the continuity of the measure  $\mu$ . □

The following Lemma is the main step in the proof of Theorem 1.1.

**Lemma 3.2.** *Let  $(X, \mathcal{B}, \mu)$  be a standard Borel probability space. Let  $T$  be an ergodic measure preserving transformation on  $X$ . Let  $\nu$  be a probability measure on  $X$  such that*

$$\frac{1}{N} \sum_{n=0}^{N-1} T_*^n \nu \xrightarrow[N \rightarrow \infty]{\text{weak-}^*} \mu$$

*Let  $f_1, \dots, f_n \in C_c(X)$ . Then, for  $\nu$ -almost every  $x \in X$ , there exists a sequence  $A \subseteq \mathbb{N}$ , of upper density 1, depending on  $x$  and the functions  $f_1, \dots, f_n$ , such that for all  $k = 1, \dots, n$ ,*

$$\lim_{\substack{N \in A \\ N \rightarrow \infty}} \frac{1}{N} \sum_{n=0}^{N-1} f_k(T^n x) = \int f_k \mu$$

Let us deduce Theorem 1.1 from this Lemma first.

**3.1. Proof of Theorem 1.1.** Let  $\mathcal{F} = \{f_k \in C_c(X) : k \in \mathbb{N}\}$  be an enumeration of a countable set of continuous functions which are dense in  $C_c(X)$  in the uniform norm. Then, it suffices to show that for  $\nu$  almost every  $x \in X$ , there exists a sequence  $A(x) \subseteq \mathbb{N}$ , of full upper density such that for all  $f \in \mathcal{F}$ ,

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(x)}} \frac{1}{N} \sum_{k=1}^N f(T^k x) = \int f d\mu \quad (3.1)$$

For each  $n \in \mathbb{N}$ , let  $\mathcal{F}_n = \{f_1, \dots, f_n\} \subset \mathcal{F}$ . By Lemma 3.2, for each  $n$ , there exists a set  $X_n$  with  $\nu(X_n) = 1$ , such that for all  $x \in X_n$ , there exists a sequence  $A(x, \mathcal{F}_n) \subseteq \mathbb{N}$ , along which the limit in (3.1) holds for all  $f \in \mathcal{F}_n$ .

Let  $Y = \bigcap_n X_n$ . Then,  $\nu(Y) = 1$ . Let  $y \in Y$ . We will build a sequence  $A(y)$  by induction from the sequences  $A(y, \mathcal{F}_n)$  via a standard argument which we will make use of several times later.

For each  $n \in \mathbb{N}$ , let  $N_n \in \mathbb{N}$  be such that for all  $N \geq N_n$  with  $N \in A(y, \mathcal{F}_n)$ , and all  $f \in \mathcal{F}_n$ ,

$$\left| \frac{1}{N} \sum_{k=1}^N f(T^k y) - \int f d\mu \right| \leq \frac{1}{n} \quad (3.2)$$

Let  $M_1 = N_1$ . If  $M_j$  has been defined, let  $M_{j+1}$  be such that the following holds

$$\begin{aligned} \frac{|A(y, \mathcal{F}_j) \cap [1, M_{j+1}]|}{M_{j+1}} &\geq 1 - \frac{1}{j} \\ \frac{M_j}{M_{j+1}} &\leq \frac{1}{j} \\ M_{j+1} &\geq N_{j+1} \end{aligned}$$

Note that the above implies that

$$\frac{|A(y, \mathcal{F}_j) \cap [M_j, M_{j+1}]|}{M_{j+1}} \geq 1 - \frac{2}{j}$$

Now, define the sequence  $A(y)$  as follows:

$$A(y) = \bigcup_{j=1}^{\infty} A(y, \mathcal{F}_j) \cap [M_j, M_{j+1}]$$

Thus, by construction, the upper density of  $A(y)$  is equal to 1. Now, let  $f \in \mathcal{F}$ . Then,  $f \in \mathcal{F}_n$  for all  $n \geq n_0$ , for some  $n_0 \in \mathbb{N}$ . Thus, for all  $N \geq M_{n_0}$  such that  $N \in A(y)$ , there exists  $j \geq n_0$ , such that  $N \in A(y, \mathcal{F}_j) \cap [M_j, M_{j+1}]$ . Thus, since  $M_j \geq N_j$ , by (3.2), the conclusion follows.

**3.2. Proof of Lemma 3.2.** For any function  $\psi$  and for every  $N \geq 1$ , let  $\mu(\psi) = \int \psi d\mu$  and let

$$A_N(\psi) = \frac{1}{N} \sum_{n=0}^{N-1} \psi \circ T^n$$

Let  $\varepsilon \in (0, 1)$ . By the usual mean ergodic theorem, for all  $k$ ,

$$A_N(f_k) \xrightarrow{L_1(\mu)} \mu(f_k)$$



Hence, we can find some  $M \gg 1$ , for all  $k \leq n$ ,

$$\int |A_M(f_k) - \mu(f_k)| d\mu < \frac{\varepsilon^3}{n}$$

Let  $\beta = \frac{\varepsilon}{C}$ , where

$$C = \max_{1 \leq k \leq n} 2\|f_k\|_{L^\infty} + 1$$

Let  $g_k = A_M(f_k) - \mu(f_k)$ . Note that  $\|g_k\|_\infty \leq C$ . For all  $k$  and for every  $N \in \mathbb{N}$ , define

$$E_{\varepsilon,N}^k = \left\{ x \in X : \sup_{1 \leq m \leq N} \left| \frac{1}{n} \sum_{l=0}^{m-1} g_k(T^l x) \right| > \varepsilon \right\}$$

Then, by the analogue of the maximal inequality, Proposition 2.1, applied to  $g_k$ , the sequence of transformations  $T_l = T^l$  and  $E_{\varepsilon,N,j} = T^{-j} E_{\varepsilon,N}^k$ , for all  $N$  sufficiently large, depending on  $\varepsilon$ , there exists  $j_{N,k} \in [0, \beta N]$  such that

$$\nu(T^{-j_{N,k}} E_{\varepsilon,N}^k) \leq \frac{12\|g_k\|_{L^1(\mu)}}{\varepsilon\beta} \leq \frac{12C\varepsilon}{n} \quad (3.3)$$

Let  $G_{N,k}^\varepsilon = X \setminus T^{-j_{N,k}} E_{\varepsilon,N}^k$  and let  $G_N^\varepsilon = \bigcap_{k=1}^n G_{N,k}^\varepsilon$ . Thus, by (3.3) and Lemma 3.1,

$$\nu \left( \limsup_{N \rightarrow \infty} G_N^\varepsilon \right) \geq 1 - 12C\varepsilon \quad (3.4)$$

Now, let  $y \in G_N^\varepsilon$  and let  $Q \in [\sqrt{\varepsilon}N, N]$ . Then, for all  $k = 1, \dots, n$ , by definition of  $E_{\varepsilon,N}^k$  and our choice of  $\beta$ ,

$$\begin{aligned} |A_{(Q+\beta N)}(g_k)(y)| &\leq \left| \frac{Q}{Q+\beta N} \frac{1}{Q} \sum_{l=j_{N,k}}^{j_{N,k}+Q-1} g_k(T^l y) \right| + \frac{\beta N \|g_k\|_{L^\infty}}{Q+\beta N} \\ &\leq |A_Q(g_k)(T^{j_{N,k}} y)| + \frac{C\beta}{\sqrt{\varepsilon}} \\ &\leq \varepsilon + \sqrt{\varepsilon} \leq 2\sqrt{\varepsilon} \end{aligned} \quad (3.5)$$

Hence, in particular, for any  $y \in \limsup_N G_N^\varepsilon$ , there exists a sequence  $N_i \rightarrow \infty$  for which (3.5) holds for all  $Q \in [\sqrt{\varepsilon}N_i, N_i]$  and for all  $k = 1, \dots, n$ . Define the following sequence for  $y \in \limsup_N G_N^\varepsilon$

$$A(y, \varepsilon) = \bigcup_{N_i: y \in G_{N_i}^\varepsilon} [(\sqrt{\varepsilon} + \beta)N_i, (1 + \beta)N_i] \cap \mathbb{N} \quad (3.6)$$

Now, a simple computation shows that for all  $N$ , and any function  $\psi$ ,

$$\begin{aligned} A_N(A_M(\psi)) &= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \psi \circ T^{n+m} \\ &= A_N(\psi) + O_M \left( \frac{\|\psi\|_{L^\infty}}{N} \right) \end{aligned} \quad (3.7)$$

Combining 3.5 and 3.7 implies that for every  $y \in \limsup_N G_N^\varepsilon$ , all  $k \leq n$  and for all  $Q \in A(y, \varepsilon)$  such that  $Q \gg M$ ,

$$|A_Q(f_k)(y) - \mu(f_k)| \leq 3\sqrt{\varepsilon} \quad (3.8)$$

Note that the choice of  $\varepsilon$  in the above considerations was arbitrary. Hence, we may consider sets  $B^{\varepsilon_i} = X \setminus \limsup_N G_N^{\varepsilon_i}$ , where  $\varepsilon_i = 1/i^2$ . Then, by (3.4) and the Borel-Cantelli lemma applied to  $B^{\varepsilon_i}$ ,

$$\nu \left( X \setminus \limsup_{i \rightarrow \infty} B^{\varepsilon_i} \right) = 1 \quad (3.9)$$

Let  $y \in X \setminus \limsup_i B^{\varepsilon_i}$ . We claim that there exists a sequence  $A(y) \subseteq \mathbb{N}$  of full upper density, such that for all  $k \leq n$ ,

$$\limsup_{\substack{N \rightarrow \infty \\ N \in A(y)}} |A_N(f_k)(y) - \mu(f_k)| = 0 \quad (3.10)$$

By (3.9), this will conclude the proof.

Since  $y \in \limsup_N G_N^{\varepsilon_i}$  for all  $i$  sufficiently large, we have sequences  $A(y, \varepsilon_i)$  as before for all  $\varepsilon_i$  sufficiently small. Without loss of generality, we may assume this holds for all  $\varepsilon_i$ . Note that by (3.6), the upper density of  $A(y, \varepsilon_i)$  is at least  $\frac{1 - \sqrt{\varepsilon_i}}{1 + \varepsilon_i/C}$ .

We build the sequence  $A(y)$  from  $A(y, \varepsilon_i)$  by induction as follows. Let  $N_i \in A(y, \varepsilon_i)$  be such that (3.8) holds for all  $k \leq n$  and all  $Q \geq N_i$ . Let  $M_1 = N_1$ . If  $M_j$  is defined, let  $M_{j+1} \in \mathbb{N}$  be such that

$$\begin{aligned} \frac{|A(y, \varepsilon_j) \cap [1, M_{j+1}]|}{M_{j+1}} &\geq \frac{1 - \sqrt{\varepsilon_j}}{1 + \varepsilon_j/C} - \frac{1}{j} \\ \frac{M_j}{M_{j+1}} &\leq \frac{1}{j} \\ M_{j+1} &\geq N_{j+1} \end{aligned}$$

This in particular, implies that

$$\frac{|A(y, \varepsilon_j) \cap [M_j, M_{j+1}]|}{M_{j+1}} \geq \frac{1 - \sqrt{\varepsilon_j}}{1 + \varepsilon_j/C} - \frac{2}{j}$$

Now, define the sequence  $A(y)$  as follows:

$$A(y) = \bigcup_{j=1}^{\infty} (A(y, \varepsilon_j) \cap [M_j, M_{j+1}])$$

Thus, since  $\varepsilon_j \rightarrow 0$ , the upper density of  $A(y)$  is equal to 1. Moreover, by (3.8) and by choice of  $M_j$ , we have that (3.10) holds as desired.

#### 4. AN ANALOGUE OF BIRKHOFF FOR SEQUENCES OF TRANSFORMATIONS - THEOREM 1.2

In this section, we prove Theorem 1.2. We will use similar ideas to those used in the proof of Theorem 1.1 by applying the weak-type maximal inequality to a carefully chosen set of continuous functions capturing the structure of the ergodic invariant measures under the transformation  $S$ .

First, we make some standard reductions. Note that since  $X$  is locally compact and second countable, by passing to the one point compactification and extending all the transformations on  $X$  trivially to the point at infinity, we may assume that  $X$  is in fact compact.

The set  $Z$  in the assumption will then be enlarged to include the point at infinity since the dirac measure at that point will be an ergodic invariant measure for  $S$ . Also, since  $X$  is

now assumed compact, the space of probability measures on it is weak-\* compact and thus we can always find limit points of infinite sequences.

*Proof of Theorem 1.2.* Let  $\varepsilon \in (0, 1)$  be fixed and let  $\varepsilon_n = \varepsilon^2/4^n$  for  $n \in \mathbb{N}$ . By regularity of the measure  $\mu$ , since  $\mu(Z) = 0$ , there exists an open set  $U_n$  containing  $Z$ , such that  $\mu(U_n \setminus K_n) < \varepsilon_n$ , for each  $n$ .

Moreover, by Urysohn's lemma, we can find a continuous function  $0 \leq f_n \leq 1$  such that  $f_n|_{K_n} \equiv 1$  and  $f_n \equiv 0$  on  $X \setminus U_n$ . Thus,

$$\|f_n\|_{L^1(\mu)} = \int f_n(x) d\mu(x) < \varepsilon_n$$

Let  $n$  be fixed. For each  $j \in \mathbb{N}$ ,  $k \leq n$  and  $\alpha \in \mathbb{R}$ , define the following set

$$E_{\alpha, N, j}^k = \left\{ x \in X : \sup_{1 \leq M \leq N} \frac{1}{M} \sum_{m=j+1}^{j+M} f_k(T_m x) > \alpha \right\}$$

Applying the analogue of the maximal inequality, Proposition 2.1, with  $f_k$ ,  $\alpha_k = \beta_k = \varepsilon_k^{1/4}$ , we get that for all  $N$  sufficiently large, depending on  $f_k$ , there exists  $j_{N, k} \in [0, \beta_k N]$  such that

$$\nu(E_{\alpha_k, N, j_{N, k}}^k) \leq \frac{12\|f_k\|_{L^1(\mu)}}{\alpha_k \beta_k} \ll \varepsilon_k^{1/2} \quad (4.1)$$

for each  $k \leq n$ .

Let  $G_{N, k} = X \setminus E_{\alpha_k, N, j_{N, k}}^k$ . Let  $y \in G_{N, k}$  and let  $Q \in [\varepsilon_k^{1/8} N, N]$ . Then, by definition of  $E_{\alpha_k, N, j_{N, k}}^k$ ,

$$\begin{aligned} \frac{1}{Q + \beta_k N} \sum_{l=1}^{Q + \beta_k N} f_k(T_l y) &\leq \frac{Q}{Q + \beta_k N} \frac{1}{Q} \sum_{l=j_{N, k} + 1}^{j_{N, k} + Q} f_k(T_l y) + \frac{\beta_k N \|f_k\|_{L^\infty}}{Q + \beta_k N} \\ &\leq \alpha_k + \frac{\beta_k}{\varepsilon_k^{1/8}} \leq 2\varepsilon_k^{1/8} \end{aligned} \quad (4.2)$$

Now, for each  $N \gg 1$ , depending on  $n$ , define the following set

$$V_{N, n} = \bigcap_{k=1}^n G_{N, k} \quad (4.3)$$

and let  $W_n = \limsup_N V_{N, n}$ . The sets  $W_n$  have the following properties:

- By (4.1), since  $\nu(G_{N, k}) \geq 1 - \varepsilon/2^k$ ,  $k = 1, \dots, n$ , we have  $\nu(V_{N, n}) \geq 1 - \varepsilon$  for all  $N \gg 1$ . In particular,

$$\nu(W_n) = \nu\left(\limsup_{N \rightarrow \infty} V_{N, n}\right) \geq 1 - \varepsilon \quad (4.4)$$

- For each  $y \in W_n$ , by (4.2) and (4.3), and noting that  $\varepsilon > \varepsilon_k$ , there exists a sequence  $A(y, n) \subseteq \mathbb{N}$  defined by

$$A(y, n) = \bigcup_{N_i: y \in V_{N_i, n}} [(\varepsilon^{1/4} + \varepsilon^{1/8})N_i, N_i] \cap \mathbb{N} \quad (4.5)$$

such that for all  $k = 1, \dots, n$  and for all  $Q \in A(y, n)$ ,

$$\frac{1}{Q} \sum_{l=1}^Q f_k(T_l y) \leq 2\varepsilon_k^{1/8} \quad (4.6)$$

Let  $W = \limsup_n W_n$ . Then, by (4.4),

$$\nu(W) \geq 1 - \varepsilon \quad (4.7)$$

Let  $y \in W$ . Then, there exists a sequence  $n_i \rightarrow \infty$ , such that  $y \in W_{n_i}$  for all  $i$ . We will construct a sequence  $A(y)$  from the sequences  $A(y, n_i)$  defined in (4.5) as follows. Let

$$\eta = \varepsilon^{1/4} + \varepsilon^{1/8}$$

First, we define a sequence  $N_i$  by induction. Let  $N_0 = 1$ . If  $N_j$  has been defined, let  $N_{j+1}$  be such that the following holds

$$\begin{aligned} \frac{|A(y, n_{j+1}) \cap [1, N_{j+1}]|}{N_{j+1}} &\geq 1 - 2\eta \\ \frac{N_j}{N_{j+1}} &\leq \eta \end{aligned}$$

This is possible since the sequences  $A(y, n)$  have upper density at least  $1 - \eta$ . These conditions imply that

$$\frac{|A(y, n_{j+1}) \cap [N_j, N_{j+1}]|}{N_{j+1}} \geq 1 - 3\eta \quad (4.8)$$

Now, define the sequence  $A(y)$  as follows:

$$A(y) = \bigcup_{j=0}^{\infty} A(y, n_{j+1}) \cap [N_j, N_{j+1}] \quad (4.9)$$

Thus, by (4.8), we get

$$\limsup_{N \rightarrow \infty} \frac{|A(y) \cap [1, N]|}{N} \geq 1 - 3\eta \quad (4.10)$$

We claim that

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(y)}} \frac{1}{N} \sum_{n=1}^N \delta_{T_n y} = \mu \quad (4.11)$$

Let  $\lambda_\infty^y$  be any weak-\* limit of the sequence  $\frac{1}{N} \sum_{n=1}^N \delta_{T_n y}$ ,  $N \in A(y)$ . First, we claim that  $\lambda_\infty^y(Z) = 0$ . Suppose otherwise. Then, by regularity of  $\lambda_\infty^y$ , there exists some compact set  $L \subseteq Z$  such that  $\lambda_\infty^y(L) > 0$ . Since  $Z = \bigcup_i K_i$  and  $K_i \subseteq K_{i+1}$  for all  $i$ , there exists some  $i_0$  such that for all  $i > i_0$ :

$$L \subseteq K_i$$

Fix some  $i > i_0$ . By definition of the functions  $f_i$ ,  $\lambda_\infty^y(f_i) \geq \lambda_\infty^y(K_i)$ . Then, for all  $n_j \geq i$  and for all  $N \in A(y, n_{j+1}) \cap [N_j, N_{j+1}] \subset A(y)$ , by (4.6), we get

$$\lambda_\infty^y(f_i) \leq 2\varepsilon_i^{1/8}$$

In particular, we get that  $\lambda_\infty^y(L) \leq 2\varepsilon_i^{1/8}$ . But, this applies to  $i > i_0$ . Thus, since  $\varepsilon_i \rightarrow 0$ , we get that  $\lambda_\infty^y(L) = 0$ , a contradiction.

Next, by our hypothesis, (after possibly intersecting  $W$  with a set of full measure),  $\lambda_\infty^y$  is  $S$ -invariant. However, all the ergodic  $S$ -invariant measures different from  $\mu$  live on  $Z$  to which  $\lambda_\infty^y$  assigns 0 mass. Thus, by the ergodic decomposition, we get that  $\lambda_\infty^y = \mu$ . Hence, the sequence  $\frac{1}{N} \sum_{n=1}^N \delta_{T_n y}$ ,  $N \in A(y)$  has  $\mu$  as its only weak-\* limit point as desired.

Thus far, we proved that for all  $y \in W$ , a set of  $\nu$  measure at least  $1 - \varepsilon$ , there exists a sequence  $A(y)$  of upper density at least  $1 - 3\eta$  such that (4.11) holds. Since  $\varepsilon$  was arbitrary, the conclusion of the theorem holds  $\nu$  almost everywhere as desired.  $\square$

## 5. INVARIANCE BY UNIPOTENT SUBGROUPS

In order to be able to apply Theorem 1.2 to prove Theorem 1.4, we shall prove that limit points of the measures considered are invariant by a one-parameter unipotent subgroup. The precise statement is Theorem 5.1 below.

Our main tool will be an adaptation of a lemma due to Chaika and Eskin [CE] in the context of the action of  $SL(2, \mathbb{R})$  on flat surfaces. We will need some facts about the structure of affine symmetric spaces before stating the precise statement.

**5.1. Structure of Affine Symmetric Spaces.** We follow the exposition in [EM] closely for the material in this section. Let  $G$  be a connected semisimple Lie group with finite center. Let  $\sigma : G \rightarrow G$  be an involution such that  $H$  is the fixed point set of  $\sigma$ . Then,  $G/H$  is an affine symmetric space. By abuse of notation, let  $\sigma$  also denote the differential of  $\sigma$  at identity. Let  $\mathfrak{g}$  denote the Lie algebra of  $G$ . Then, we have

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$$

where  $\mathfrak{h}$  is the eigenspace corresponding to the eigenvalue 1 of  $\sigma$  and  $\mathfrak{p}$  corresponds to the  $-1$  eigenspace, and  $\mathfrak{h}$  is the Lie algebra of  $H$ .

It is well known (Proposition 7.1.1, [Sch]) that one can find a Cartan involution  $\theta$  of  $G$  commuting with  $\sigma$ . Let  $\theta$  also denote its differential at identity. Then, similarly  $\mathfrak{g}$  splits as

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{q}$$

where  $\mathfrak{k}$  (resp.  $\mathfrak{q}$ ) is the  $+1$  (resp.  $-1$ ) eigenspace of  $\theta$ . Since  $\theta$  is a Cartan involution,  $\mathfrak{k}$  is the Lie algebra of a maximal compact subgroup, denote it by  $K$ .

Now, let  $\mathfrak{a}$  be a maximal abelian subspace of  $\mathfrak{p} \cap \mathfrak{q}$ . Then,  $\mathfrak{a}$  is the Lie algebra of a maximal abelian subgroup  $A$  and the exponential map  $\mathfrak{a} \rightarrow A$  is a diffeomorphism.

Recall that  $G$  admits a decomposition of the form  $G = KAH$  (See [Sch], Proposition 7.1.3 or [EM], Proposition 4.2). Elements of the fiber of the map  $(k, a, h) \mapsto kah$  have the form  $(kl, a, l^{-1}h)$  for some element  $l \in K \cap H$ . In particular, the fiber lies in a compact group.

Consider the adjoint action of  $\mathfrak{a}$  on  $\mathfrak{g}$ . There exists a finite subset  $\Sigma \subset \mathfrak{a}^*$  of non-zero elements of the dual of  $\mathfrak{a}$  such that  $\mathfrak{g}$  splits as

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Sigma} \mathfrak{g}_\alpha$$

such that for all  $X \in \mathfrak{a}$  and all  $Z \in \mathfrak{g}_\alpha$ ,

$$ad_X(Z) = \alpha(X)Z$$

And, for  $Z \in \mathfrak{g}_0$ ,  $ad_X(Z) = 0$ . Recall that the subspaces

$$\{X \in \mathfrak{a} : \alpha(X) = 0\}$$

for  $\alpha \in \Sigma$  divide  $\mathfrak{a}$  into a finite collection of cones, called Weyl Chambers.

Let  $\mathfrak{C}$  be one such Weyl chamber. Let  $\Sigma^+$  denote the set of  $\alpha \in \Sigma$  such that  $\alpha(X) > 0$  for all  $X \in \mathfrak{C}$ . We call  $\Sigma^+$  the set of positive roots relative to  $\mathfrak{C}$ . Then,  $\mathfrak{g}$  splits as follows:

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{g}_0 \oplus \mathfrak{n}^+$$

where  $\mathfrak{n}^- = \bigoplus_{\alpha \in \Sigma - \Sigma^+} \mathfrak{g}_\alpha$  and  $\mathfrak{n}^+ = \bigoplus_{\alpha \in \Sigma^+} \mathfrak{g}_\alpha$ .

**5.2. Unipotent Invariance.** We need to fix some notation before stating the main theorem. For  $g \in G$ , let  $Ad(g)$  denote the associated linear map on  $Lie(G)$  via the Adjoint action.

Fix some norm on  $Lie(G)$  inducing the metric on  $G$  and denote it by  $\|\cdot\|$ . This norm induces a matrix norm for the Adjoint maps. Let  $\|Ad(g)\|$  denote such matrix norm.

The following is the main theorem of this section. See remark after Theorem 6.1 below for a discussion of the growth condition required in the theorem.

**Theorem 5.1.** *In the notation above, if  $g_n \in G$  is a sequence tending to infinity in  $G/H$  and satisfying the following growth condition:*

(1) *There exist constants  $C > 1$ ,  $\lambda > 0$  such that for all  $n \gg 1$ ,*

$$\frac{1}{C}e^{\lambda n} \leq \|Ad(g_n)\| \leq Ce^{\lambda n}$$

(2) *Writing  $g_n = k_n a_n h_n$ , we have that  $\|Ad(h_n^{-1})\|$  is uniformly bounded for all  $n$ .*

*Then, after passing to a subsequence of  $g_n$ , for  $\mu_H$ -almost every  $x \in G/\Gamma$ , any weak-\* limit point of the sequence of measures*

$$\left\{ \frac{1}{N} \sum_{n=1}^N \delta_{g_n x} \right\}$$

*is invariant by a one-parameter unipotent subgroup of  $G$ .*

**Remark 5.2.** The second growth condition in Theorem 5.1 makes sense, since the element  $h_n$  in the decomposition of  $g_n$  is unique up to left multiplication by elements inside the compact group  $H \cap K$ .

**5.3. Expansion Properties of the Adjoint Action.** We shall need the following lemma regarding the Adjoint action of  $G$ . This lemma uses a standard argument exploiting the relationship between diagonalizable elements and their associated horospherical subgroups. We also make use of the structure of affine symmetric spaces.

**Lemma 5.3.** *Let  $G$ ,  $H$  and  $g_n$  be as in Theorem 5.1. Then, there exists a sequence  $v_n \rightarrow 0 \in Lie(H)$  satisfying the following for all  $n$ ,*

$$\|v_n\| \ll \frac{1}{\|Ad(g_n)\|}$$

*and such that after passing to a subsequence of the  $g_n$ 's, we have*

$$g_n \exp(v_n) g_n^{-1} \rightarrow u \neq id$$

*where  $u$  is an Ad-unipotent element in  $G$ .*

**Example 5.4.** Let  $G = SL(2, \mathbb{R})$  and  $H = K = SO(2)$ . Let  $t_n \rightarrow +\infty$  be a sequence. Let  $g_n$  be the following sequence:

$$g_n = \begin{pmatrix} 1 & t_n \\ 0 & 1 \end{pmatrix}$$

Let  $k_\theta \in K$ . Then,

$$g_n k_\theta g_n^{-1} = \begin{pmatrix} \cos(\theta) - t_n \sin(\theta) & (t_n^2 + 1) \sin(\theta) \\ -\sin(\theta) & \cos(\theta) + t_n \sin(\theta) \end{pmatrix}$$

Let  $\alpha \in (0, 1)$  be a fixed real number. For all large  $n$ , let  $\theta_n$  be such that  $t_n^2 \sin(\theta_n) = \alpha$ . Then, as  $n \rightarrow \infty$ ,  $\theta_n \rightarrow 0$  and  $t_n \sin(\theta_n) \rightarrow 0$ . Hence, we get

$$g_n k_{\theta_n} g_n^{-1} \rightarrow u(\alpha) = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \neq id$$

We will need the following fact for the proof of Lemma 5.3.

**Lemma 5.5.** (*Lemma 3, [Moz]*) *If  $G$  is semisimple over  $\mathbb{R}$  with finite center, then the Adjoint representation  $Ad : G \rightarrow GL(\mathfrak{g})$  is a proper map.*

We are now ready for the proof.

5.3.1. *Proof of Lemma 5.3.* Write  $g_n = k_n a_n h_n$ . Then, by passing to a subsequence of  $g_n$ , we may assume that there exists a single Weyl chamber  $\mathfrak{C}$  such that  $a_n = \exp(X_n)$  and  $X_n \in \mathfrak{C}$  for all  $n$ . Let  $\Sigma^+$  be a set of positive roots associated with  $\mathfrak{C}$ .

First, we'll assume that  $g_n \in KA$  and write  $g_n = k_n a_n$ . Note that the map  $Ad : G \rightarrow GL(\mathfrak{g})$  is proper by Lemma 5.5. In particular, by assumption, since  $g_n \rightarrow \infty$ , we have

$$\|Ad(g_n)\| \rightarrow \infty$$

We claim that the image of  $\mathfrak{h}$  under the projection onto  $\mathfrak{n}^+$  is non-zero. To see this, note that given  $X, Y \in \mathfrak{p}$ , we have that

$$\sigma(ad_X(Y)) = ad_{\sigma(X)}(\sigma(Y)) = ad_X(Y)$$

Hence,  $ad_X(Y) \in \mathfrak{h}$ . On the other hand, if  $X \in \mathfrak{a} \subseteq \mathfrak{p}$  and  $Y \in \mathfrak{g}_\alpha$  for some  $\alpha \neq 0 \in \Sigma$ , then  $ad_X(Y) \in \mathfrak{g}_\alpha$ . In particular, this implies

$$\mathfrak{n}^+ \cap \mathfrak{p} = \{0\}$$

Thus, given any  $X \neq 0 \in \mathfrak{n}^+$ , the element  $X + \sigma(X) \neq 0$  and is  $\sigma$  invariant and hence belongs to  $\mathfrak{h}$ .

Next, note that since  $\sigma(X) = -X$  for all  $X \in \mathfrak{a}$ , we have  $\sigma(\mathfrak{n}^+) = \mathfrak{n}^-$ . Thus, in particular, for any  $v \in \mathfrak{n}^+$ ,

$$\|Ad(g_n)(v)\| \rightarrow \infty, Ad(g_n)(\sigma(v)) \rightarrow 0 \quad (5.1)$$

Let  $V = \{v_\alpha \neq 0 \in \mathfrak{g}_\alpha : \alpha \in \Sigma^+\}$  be fixed. For each  $n$ , let  $v_{\alpha_n} \in V$  be such that

$$\alpha_n(X_n) = \max \{\alpha(X_n) : \alpha \in \Sigma^+\}$$

where  $X_n \in \mathfrak{C}$  was such that  $a_n = \exp(X_n)$ . Now, for each  $n$ , let

$$v_n = \frac{v_{\alpha_n} + \sigma(v_{\alpha_n})}{\|Ad(g_n)\|} \quad (5.2)$$

Then, for all  $n$ ,  $v_n \neq 0$  in  $\mathfrak{h}$  and satisfies

$$\|v_n\| \ll \frac{1}{\|Ad(g_n)\|}$$

Moreover, by the standard identity  $Ad(\exp) = \exp(ad)$  and by 5.1,

$$Ad(g_n)(v_n) = \frac{e^{\alpha_n(X_n)}}{\|Ad(g_n)\|} Ad(k_n)(v_{\alpha_n}) + O(1)$$

By compactness of  $K$ , we have  $\|Ad(g_n)\| \ll \|Ad(a_n)\|$ . But, by our choice of  $\alpha_n$ ,  $e^{\alpha_n(X_n)}$  is the largest eigenvalue of  $Ad(a_n)$  and  $Ad(a_n)$  is diagonalizable. Thus,  $\|Ad(a_n)\| \ll e^{\alpha_n(X_n)}$  and so we get

$$\|v_{\alpha_n}\| \ll \|Ad(g_n)(v_n)\| \leq \|Ad(g_n)\| \|v_n\| \ll \|v_{\alpha_n}\|$$

Hence, by passing to a subsequence, we get that

$$g_n \exp(v_n) g_n^{-1} = \exp(Ad(g_n)(v_n)) \rightarrow u \neq id$$

Since  $v_n \rightarrow 0$ , we have that  $\exp(v_n) \rightarrow id$ . Hence, all the eigenvalues of  $Ad(\exp(v_n))$  converge to 1. Since conjugation doesn't change eigenvalues, we get that  $u$  must be an  $Ad$ -unipotent element, which finishes the proof in the case  $g_n \in KA$ .

For the general case, by  $KAH$  decomposition, we write  $g_n = k_n a_n h_n$ . Then, we can find  $v_n \in \mathfrak{h}$  as above such that  $Ad(k_n a_n)(\exp(v_n)) \rightarrow u \neq id$ . Thus, the elements  $w_n = h_n^{-1} v_n h_n$  will satisfy  $Ad(g_n)(\exp(w_n)) = Ad(k_n a_n)(\exp(v_n)) \rightarrow u$ . Taking  $h_n = \exp(w_n)$ .

By our assumption on the boundedness of  $\|Ad(h_n^{-1})\|$ , we get

$$\|w_n\| \leq \|Ad(h_n^{-1})\| \|v_n\| \ll \frac{1}{\|Ad(g_n)\|}$$

which concludes the proof.

**5.4. Decay of Correlations.** Let  $\Gamma$ ,  $H$  and  $G$  be as above and let  $g_n$  be a sequence tending to infinity in  $G/H$ . Let  $\varphi \in C_c^\infty(G/\Gamma)$ . Pass to a subsequence such that the conclusion of Lemma 5.3 holds. For each  $n$ , define the following function on  $H\Gamma/\Gamma$ :

$$f_n(h\Gamma) = \varphi(g_n \exp(v_n) h\Gamma) - \varphi(g_n h\Gamma) \quad (5.3)$$

where  $v_n \in Lie(H)$  is as in the statement of Lemma 5.3.

The reason for defining such functions is the following

**Proposition 5.6.** *To prove Theorem 5.1, it suffices to show that for  $\mu_H$  almost every  $x \in H\Gamma/\Gamma$ , the following holds:*

$$\frac{1}{N} \sum_{n=1}^N f_n(x) \rightarrow 0$$

*Proof.* Let  $h_n = \exp(v_n)$ . By Lemma 5.3, we have  $g_n h_n g_n^{-1} \rightarrow u$ , where  $u$  is a non-trivial  $Ad$ -unipotent element. Since,  $\varphi$  is uniformly Lipschitz, we have

$$\varphi(ug_n x) - \varphi(g_n h_n x) = \varphi(ug_n x) - \varphi(g_n h_n g_n^{-1} g_n x) = O(d(u, g_n h_n g_n^{-1}))$$

where  $d(.,.)$  is the right invariant metric on  $G$ . Hence,  $|\varphi(ug_n x) - \varphi(g_n h_n x)| \rightarrow 0$  as  $n \rightarrow \infty$ .

Hence, by assumption,

$$\left| \frac{1}{N} \sum_{n=1}^N (\varphi(ug_n x) - \varphi(g_n x)) \right| \leq \frac{1}{N} \sum_{n=1}^N |\varphi(ug_n x) - \varphi(g_n h_n x)| + \left| \frac{1}{N} \sum_{n=1}^N f_n(x) \right| \rightarrow 0$$

But,  $\varphi$  was an arbitrary function. Thus, any limit point must be invariant by the group generated by  $u$ .



□

In order to use Proposition 5.6, we shall need the following estimate on the correlation between the functions  $f_n$ . This lemma is an adaptation of the key technical lemma in [CE] (Lemma 3.3) to our setting. we shall need the following definition.

**Definition 5.7.** For any  $x \in H\Gamma/\Gamma \cong H/(H \cap \Gamma)$ , the injectivity radius at  $x$ , denoted by  $inj_x$  is defined to be the infimum over all  $r > 0$  such that the map  $h \mapsto hx$  is injective on the ball of radius  $r$  around identity in  $H$ .

**Lemma 5.8.** For all  $n \geq m \gg 1$  such that  $\|Ad(g_m)\|/\|Ad(g_n)\|$  is sufficiently small, the following holds

$$\int f_n(h\Gamma) f_m(h\Gamma) d\mu_H(h\Gamma) = O\left(\left(\frac{\|Ad(g_m)\|}{\|Ad(g_n)\|}\right)^{1/2}\right) \quad (5.4)$$

*Proof.* Let  $v_n$  be as in the definition of the functions  $f_n$  and let  $h_n = \exp(v_n)$ . Let  $d_n = \|Ad(g_n)\|$  and  $d_m = \|Ad(g_m)\|$ . Define

$$r = \left(\frac{1}{d_m d_n}\right)^{1/2} \quad (5.5)$$

Let  $B_H(e, r)$  denote the ball of radius  $r$  around the identity in  $H$ . By abuse of notation, we'll use  $\mu_H$  to denote the Haar measure on  $H$  and on  $H\Gamma/\Gamma$ .

Let  $\psi : H\Gamma/\Gamma \rightarrow \mathbb{R}$  be any integrable function. Then, by Fubini's theorem and left  $H$ -invariance of  $\mu_H$ ,

$$\int_{H\Gamma/\Gamma} \psi(x) d\mu_H(x) = \int_{H\Gamma/\Gamma} \frac{1}{\mu_H(B_H(e, r))} \int_{B_H(e, r)} \psi(hx) d\mu_H(h) d\mu_H(x)$$

Hence, since  $\mu_H$  is a probability measure on  $H\Gamma/\Gamma$ , it suffices to prove for all  $x \in H\Gamma/\Gamma$ ,

$$\frac{1}{\mu_H(B_H(e, r))} \int_{B_H(e, r)} f_n(hx) f_m(hx) d\mu_H(h) = O\left(\left(\frac{\|Ad(g_m)\|}{\|Ad(g_n)\|}\right)^{1/2}\right) \quad (5.6)$$

Define the following set

$$Thick_r = \{x \in H\Gamma/\Gamma : inj_x \geq r\}$$

where  $inj_x$  denotes the injectivity radius at  $x$  in  $H/(H \cap \Gamma)$ . Let  $Thin_r = H\Gamma/\Gamma - Thick_r$ . Let  $w \in Thick_r$  be fixed.

Let  $B_r$  denote the ball of radius  $r$  around the  $w$  in  $H\Gamma/\Gamma$  in the metric induced by the metric on  $G$ . Then, for every  $x \in B_r$ , there exists some  $l \in B_H(e, r)$  such that  $x = lw$ . Since  $\varphi \in C_c^\infty(G/\Gamma)$ ,  $\varphi$  is uniformly Lipschitz. Thus, we get

$$\varphi(g_m h_m x) - \varphi(g_m h_m w) = \varphi(g_m h_m l w) - \varphi(g_m h_m w) = O(d(g_m h_m l h_m^{-1} g_m^{-1}, e))$$

Since the sequences  $d_n, d_m$  are tending to infinity, for all  $n, m$  sufficiently large,  $r$  will be small enough so that the exponential map is a diffeomorphism from a neighborhood of 0 in  $\mathfrak{h} = Lie(H)$  onto  $B_H(e, r)$ .

Thus, we can write  $l = \exp(v)$  for some  $v \in \mathfrak{h}$ . So, we have

$$\|Ad(g_m h_m)(v)\| \leq \|Ad(g_m)\| \cdot \|Ad(h_m)\| \cdot \|v\|$$

where for any  $g \in G$ .

But, since  $h_m \rightarrow id$  as  $m \rightarrow \infty$  and since the norm is continuous, for all  $m$  sufficiently large, we have  $\|Ad(h_m)\| \ll 1$ .

Moreover, since the differential of the exponential map at 0 is the identity, its Jacobian is 1 at 0 and hence, when  $r$  is sufficiently small, we have  $\|v\| \ll d(l, e) \leq r$ . Combining these estimates, we get for all  $x \in B_r$ ,

$$\|Ad(g_m h_m)(v)\| = O(\|Ad(g_m)\| r) = O\left(\left(\frac{\|Ad(g_m)\|}{\|Ad(g_n)\|}\right)^{1/2}\right)$$

But, as before, the exponential map is nearly an isometry near identity. Hence, when  $\|Ad(g_n)\|$  is sufficiently larger than  $\|Ad(g_m)\|$ ,  $Ad(g_m h_m)(v)$  will be sufficiently close to 0 so that  $d(\exp(Ad(g_m h_m)(v)), e) \ll \|Ad(g_m h_m)(v)\|$  up to absolute constants. Thus, we get for all  $x \in B_r$ ,

$$\varphi(g_m h_m x) - \varphi(g_m h_m w) = O\left(\left(\frac{\|Ad(g_m)\|}{\|Ad(g_n)\|}\right)^{1/2}\right)$$

Similarly, we get the same estimate for  $\varphi(g_m x) - \varphi(g_m w)$  for all  $x \in B_r$ . Thus, by definition of  $f_m$ , we get

$$f_m(x) - f_m(w) = O(\|Ad(g_m)\| r)$$

Thus, we get that

$$\frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x) f_m(x) d\mu_H(x) = \frac{f_m(w)}{\mu_H(B_r)} \int_{B_r} f_n(x) d\mu_H(x) + O(\|Ad(g_m)\| r) \quad (5.7)$$

Next, note that by definition of  $f_n$  and left-invariance of  $\mu_H$ ,

$$\begin{aligned} \frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x) d\mu_H(x) &= \frac{1}{\mu_H(B_r)} \int_{h_n B_r} \varphi(g_n x) d\mu_H(x) - \frac{1}{\mu_H(B_r)} \int_{B_r} \varphi(g_n x) d\mu_H(x) \\ &= O\left(\frac{\mu_H(h_n B_r \triangle B_r)}{\mu_H(B_r)}\right) \end{aligned}$$

Since  $w \in Thick_r$ ,  $B_r$  isometric to  $B_H(e, r)$ . Hence, for all  $r$  sufficiently small, we may apply Proposition 5.9 below to get

$$\frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x) d\mu_H(x) = O\left(\frac{d(h_n, e)}{r}\right) \quad (5.8)$$

Note that we are implicitly assuming that  $h_n \in B_H(e, r)$ . To see that this is the case, observe that for  $n$  sufficiently large,  $h_n = \exp(v_n)$  will be sufficiently close to identity and hence, we have

$$d(h_n, e) \ll \|v_n\|$$

But, by Lemma 5.3, we have  $\|v_n\| \ll 1/\|Ad(g_n)\|$  for all  $n$  sufficiently large. But, since  $\|Ad(g_n)\| \geq \|Ad(g_m)\|$  by assumption, we have  $\|Ad(g_n)\| \geq 1/r$ . Thus, in particular,  $h_n$  will be contained in a ball of radius comparable to  $r$  for all large  $n$ , which doesn't affect our estimate.

Moreover, this observation, along with (5.8), imply that

$$\frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x) d\mu_H(x) = O\left(\frac{\|Ad(g_n)\|}{r}\right) \quad (5.9)$$

Combining this estimate with (5.6) and (5.7) gives

$$\int_{Thick_r} f_n(h\Gamma) f_m(h\Gamma) d\mu_H(h\Gamma) = O\left(\left(\frac{\|Ad(g_m)\|}{\|Ad(g_n)\|}\right)^{1/2}\right) \quad (5.10)$$

Finally, since  $\mu_H(Thin_r) \ll r^k$ , for some  $k \in \mathbb{N}$  as  $r \rightarrow 0$  by ??, the conclusion of the lemma follows.  $\square$

5.4.1. *A measure estimate.* The following estimate was used in the proof of Lemma 5.8.

**Proposition 5.9.** *Let  $H$  be a Lie group and let  $B_r$  denote a ball of radius  $r > 0$  around the identity in  $H$ . Then, for all  $r > 0$  sufficiently small and all  $h \in B_r$ ,*

$$\frac{\mu_H(hB_r \triangle B_r)}{\mu_H(B_r)} = O\left(\frac{d(h, e)}{r}\right)$$

where  $\mu_H$  denotes a left-invariant Haar measure on  $H$  and  $d(., .)$  denotes a right invariant metric.

*Proof.* Let  $\mathfrak{h} = Lie(H)$ . Fix a norm on  $\mathfrak{h}$  inducing the metric  $d$ . Let  $r > 0$  be small enough such that the exponential map is a diffeomorphism from a ball around 0 in  $\mathfrak{h}$  onto  $B_r$ . Since the differential of the exponential is the identity at 0, such ball will have a radius comparable to  $r$ , denote it by  $B_{r'}^{\mathfrak{h}}$ .

Let  $g \in B_r$ . Let  $X, Y \in \mathfrak{h}$  be such that  $h = \exp(X)$  and  $g = \exp(Y)$ . Then, if  $r$  is sufficiently small, by the Campell-Baker-Hausdorff formula, there exists some  $Z \in \mathfrak{h}$  so that  $hg = \exp(Z)$  and

$$Z - Y = X + o(\|X\|)$$

In particular, there is some  $C \geq 1$  such that  $hB_r \subseteq \exp(B_{r'}^{\mathfrak{h}} + CX)$ . And, hence, we get that

$$hB_r \triangle B_r \subseteq \exp((B_{r'}^{\mathfrak{h}} + CX) \triangle B_{r'}^{\mathfrak{h}})$$

Let  $Leb$  denote the Lebesgue measure on  $\mathfrak{h}$ . It is then a standard fact from convex euclidean geometry that

$$Leb((B_{r'}^{\mathfrak{h}} + CX) \triangle B_{r'}^{\mathfrak{h}}) \ll \|X\| r^{dimH-1} \quad (5.11)$$

where the implicit constants are absolute and depend only on the dimension (see for example [Gro]). Here we are using that a ball in the norm on  $\mathfrak{h}$  is equivalent to a standard euclidean ball of comparable radius.

Again, since the differential of the exponential is the identity at 0, the Haar measure on  $H$  near identity is comparable up to absolute constants with the pushforward of the Lebesgue measure under the exponential map.

In particular, one has  $\mu_H(B_r) \ll r^{dimH}$ . Combining this with (5.11) gives the desired conclusion.  $\square$

**5.5. Law of Large Numbers.** This section is dedicated to the proof of Proposition 5.10 below. The proof follows the same ideas as in the proof of the strong law of large numbers and is very similar to the proof in [CE]. The main difference is the need to handle the general growth condition in Theorem 5.1.

Recall the definition of the functions  $f_n$  from the previous section. By Proposition 5.6, the following proposition concludes the proof of Theorem 5.1.

**Proposition 5.10.** *Under the same hypotheses of Theorem 5.1, for  $\mu_H$  almost every  $x \in H\Gamma/\Gamma$ ,*

$$\frac{1}{N} \sum_{n=1}^N f_n(x) \rightarrow 0$$

*Proof.* For  $x \in H\Gamma/\Gamma$  and  $N \in \mathbb{N}$ , let  $S_N(f)(x) = \sum_{n=1}^N f_n(x)$ . By assumption, there exists  $\lambda > 0$  such that for all  $n \geq m \gg 1$

$$\frac{||Ad(g_m)||}{||Ad(g_n)||} \ll e^{\lambda(m-n)} \quad (5.12)$$

Then, we have

$$\begin{aligned} \int |S_N(f)(x)|^2 d\mu_H(x) &= \sum_{1 \leq n, m \leq N} \int f_n(x) f_m(x) d\mu_H(x) \\ &= O(N^{3/2}) + \sum_{|n-m| \geq N^{1/2}} \int f_n(x) f_m(x) d\mu_H(x) \end{aligned}$$

Here we estimated the number of pairs  $(m, n)$  with  $|m - n| < N^{1/2}$  using the area between the 2 lines  $m \pm n = N^{1/2}$  in the square  $[0, N]^2$ .

But, by (5.12), when  $N \gg 1$ , for  $n \geq m$  such that  $|n - m| \geq N^{1/2}$ , we have that  $||Ad(g_m)||/||Ad(g_n)||$  will be sufficiently large so that Lemma 5.8 applies. This implies that for all  $N \gg 1$ :

$$\begin{aligned} \frac{1}{N^2} \int |S_N(f)(x)|^2 d\mu_H(x) &= O(N^{-1/2}) + O\left(e^{\frac{-\lambda N^{1/2}}{2}}\right) \\ &= O(N^{-1/2}) \end{aligned}$$

Let  $\varepsilon > 0$ . Then, by the Chebyshev-Markov inequality,

$$\mu_H \left( \left\{ x : \left| \frac{S_N(f)(x)}{N} \right| > \varepsilon \right\} \right) \ll \frac{N^{-1/2}}{\varepsilon^2}$$

For all  $k \in \mathbb{N}$ , let  $N_k = k^4$ . Thus, the above observation shows that the sequence  $N_k^{-1/2}$  is summable. Hence, by the Borel-Cantelli lemma, we have

$$\mu_H \left( x : \left| \frac{S_{N_k}(f)(x)}{N_k} \right| > \varepsilon \text{ for infinitely many } k \right) = 0$$

Since  $\varepsilon$  was arbitrary, by taking a countable sequence  $\varepsilon_i$  decreasing to 0, we conclude that for  $\mu_H$  almost every  $x$ ,

$$\lim_{k \rightarrow \infty} \frac{S_{N_k}(f)(x)}{N_k} = 0$$

We are left with bootstrapping this conclusion to all sequences, for which we use a standard interpolation argument. Let  $M_i \rightarrow \infty$  be a sequence. Observe that for each  $M_i \in \mathbb{N}$ , there exists some  $k_i \in \mathbb{N}$  such that  $N_{k_i} \leq M_i \leq N_{k_i+1}$ .

Moreover, we have  $N_{k_i+1} - N_{k_i} = O(k_i^3)$ . Thus, we get that

$$\left| \frac{S_{M_i}(f)(x)}{M_i} \right| \leq \frac{N_{k_i}}{M_i} \left| \frac{S_{N_{k_i}}(f)(x)}{N_{k_i}} \right| + O(k_i^{-1}) \xrightarrow{i \rightarrow \infty} 0$$

as desired.  $\square$

## 6. PROOF OF THEOREM 1.4

The more precise statement of Theorem 1.4 is stated and proved in this section. Let the notation be the same as in § 5.

**Theorem 6.1.** *Let  $g_n$  be a sequence tending to infinity in  $G/H$  and satisfying the following growth condition:*

(1) *There exist constants  $C > 1$ ,  $\lambda > 0$  such that for all  $n \gg 1$ ,*

$$\frac{1}{C}e^{\lambda n} \leq \|Ad(g_n)\| \leq Ce^{\lambda n}$$

(2) *Writing  $g_n = k_n a_n h_n$ , we have that  $\|Ad(h_n^{-1})\|$  is uniformly bounded for all  $n$ .*

*Then, after passing to a subsequence of  $g_n$ , for  $\mu_{H\Gamma/\Gamma}$  almost every  $x \in G/\Gamma$ , there exists a sequence  $A(x) \subseteq \mathbb{N}$  of upper density 1 such that*

$$\lim_{\substack{N \rightarrow \infty \\ N \in A(x)}} \frac{1}{N} \sum_{n=1}^N \delta_{g_n x} = \mu_{G/\Gamma}$$

*where  $\mu_{G/\Gamma}$  (resp.  $\mu_{H\Gamma/\Gamma}$ ) denote the  $G$ -invariant (resp.  $H$ -invariant) Haar probability measure on  $G/\Gamma$  (resp. the closed orbit  $H\Gamma/\Gamma \cong H/(H \cap \Gamma)$ ).*

**Remark 6.2.** Note that the growth rate of  $\|Ad(g_n)\|$  required by this theorem is not the most general one which works with our techniques. It is possible to obtain the same conclusions assuming there exist constants  $\lambda, \varepsilon > 0$  such that  $e^{\lambda n^\varepsilon} \ll \|Ad(g_n)\| \ll e^{\lambda n^\varepsilon}$ .

*Proof of Theorem 6.1.* First, pass to a subsequence of  $g_n$  such that the conclusion of Theorem 5.1 holds and denote this subsequence by  $g_n$  as well.

We will apply Theorem 1.2 with  $X = G/\Gamma$ ,  $\mu = \mu_{G/\Gamma}$ ,  $\nu = \mu_H$  and  $T_n = g_n$ . Then, Theorem 1.2 in [EM] implies that

$$(T_n)_* \nu \rightarrow \mu$$

In particular, this implies condition (1) of Theorem 1.2. Moreover, by Theorem 5.1, condition (2) is also satisfied.

Let  $\mathcal{L}$  denote the collection of proper analytic subgroups  $L$  of  $G$  such that  $L \cap \Gamma$  is a lattice. Then,  $\mathcal{L}$  is a countable set [Rat]. Let  $U$  denote the one-parameter unipotent subgroup in the conclusion of Theorem 5.1. For  $L \in \mathcal{L}$ , define the following set

$$N(L, U) = \{g \in G : g^{-1}Ug \subseteq L\}$$

Let  $\pi : G \rightarrow G/\Gamma$  denote the natural projection. The set  $Z$  appearing in the hypotheses of Theorem 1.2 will be defined to be

$$Z = \bigcup_{L \in \mathcal{L}} \pi(N(L, U))$$

Then, since  $Z$  is a countable union of analytic subvarieties of  $G/\Gamma$  [Rat],  $Z$  admits a filtration by compact sets. Moreover, since  $\mathcal{L}$  is countable, and  $\mu_{G/\Gamma}(\pi(N(L, U))) = 0$  for all  $L \in \mathcal{L}$ , we have

$$\mu_{G/\Gamma}(Z) = 0$$

Finally, by Ratner's measure rigidity theorem [Rat], any ergodic  $U$  invariant probability measure  $\lambda \neq \mu_{G/\Gamma}$  is supported on  $N(L, U)$  (in fact supported on a single closed orbit of a

conjugate of  $L$ ) for some  $L \in \mathcal{L}$ . Thus, all the hypotheses of Theorem 1.2 are verified and hence the conclusion of Theorem 1.4 follows.  $\square$

**6.1. Proof of Corollary 1.5.** The sequence  $u(t_n)$  is tending to  $\infty$  in  $G = SL(2, \mathbb{R})$  and hence it tends to infinity in  $G/K$ , by compactness of  $K$ . It is also straightforward to check that  $\|Ad(g_n)\| \ll e^{2\lambda n}$  by calculating  $Ad(g_n)$  in the standard basis for the Lie algebra of  $SL(2, \mathbb{R})$ . Moreover, since  $K$  is fixed by a Cartan involution, it's in particular a symmetric group. The compactness of  $K$  implies the second growth condition in Theorem 6.1.

Next, the argument in Example 5.4 shows that we in fact don't need to pass to a subsequence of  $g_n$  in this case to get invariance by unipotent. Thus, the first claim of the corollary follows by Theorem 6.1.

When  $G/\Gamma$  is compact, we use Theorem 5.1 which implies that for almost every  $\theta$ , any limit point of the measures  $\frac{1}{N} \sum_{n=1}^N \delta_{u(t_n)k_\theta x}$  are invariant by a unipotent element. Moreover, the compactness of  $G/\Gamma$  insures that non-zero limit points of these measures always exist.

But, the action by unipotent elements on  $G/\Gamma$  is uniquely ergodic in this case. Thus, the conclusion follows.

## REFERENCES

- [Bou] Jean Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publications Mathématiques de l'IHS* **69**(1989), 5–41.
- [CE] Jon Chaika and Alex Eskin. Every flat surface is birkhoff and oseledets generic in almost every direction. *Journal of Modern Dynamics* **9**(2015), 1–23.
- [EW] Manfred Einsiedler and Thomas Ward. *Ergodic Theory*. Springer Nature, 2011.
- [EM] Alex Eskin and Curt McMullen. Mixing, counting, and equidistribution in Lie groups. *Duke Math. J.* **71**(07 1993), 181–209.
- [FFT] Livio Flaminio, Giovanni Forni, and James Tanis. Effective equidistribution of twisted horocycle flows and horocycle maps. *Geometric and Functional Analysis* **26**(2016), 1359–1448.
- [Gro] H. Groemer. On the symmetric difference metric for convex bodies. *Beitrage zur Algebra und Geometrie* **41**(2000), 107–114.
- [Kat] Asaf Katz. On mixing and sparse ergodic theorems. *arXiv preprint arXiv:1612.01146* (2016).
- [KSW] Dmitry Kleinbock, Ronggang Shi, and Barak Weiss. Pointwise equidistribution with an error rate and with respect to unbounded functions. *Mathematische Annalen* **367**(2017), 857–879.
- [Moz] Shahar Mozes. Mixing of all orders of Lie groups actions. *Inventiones mathematicae* **107**(1992), 235–242.
- [Rat] Marina Ratner. On Raghunathan's Measure Conjecture. *Annals of Mathematics* **134**(1991), 545–607.
- [SU] Peter Sarnak and Adrián Ubis. The horocycle flow at prime times. *Journal de Mathématiques Pures et Appliquées* **103**(feb 2015), 575–618.
- [Sch] Henrik Schlichtkrull. *Hyperfunctions and Harmonic Analysis on Symmetric Spaces (Progress in Mathematics)*. Birkhuser, 1984.
- [SY] Nimish Shah and Lei Yang. Equidistribution of curves in homogeneous spaces and Dirichlet's approximation theorem for matrices, 2016.
- [Sha1] Nimish A. Shah. Asymptotic evolution of smooth curves under geodesic flow on hyperbolic manifolds. *Duke Math. J.* **148**(06 2009), 281–304.
- [Sha2] Nimish A. Shah. Equidistribution of expanding translates of curves and Dirichlet's theorem on diophantine approximation. *Inventiones mathematicae* **177**(2009), 509–532.
- [Sha3] Nimish A Shah. Expanding translates of curves and Dirichlet-Minkowski theorem on linear forms. *Journal of the American Mathematical Society* **23**(2010), 563–589.
- [Shi] Ronggang Shi. Pointwise equidistribution for one parameter diagonalizable group action on homogeneous space, 2014.

- [TV] James Tanis and Pankaj Vishe. Uniform Bounds for Period Integrals and Sparse Equidistribution. *International Mathematics Research Notices* **2015**(2015), 13728.
- [Ven] A. Venkatesh. Sparse equidistribution problems, period bounds and subconvexity. *Annals of Math.* **172**(2010), 989–1094.
- [Zhe] Cheng Zheng. Sparse equidistribution of unipotent orbits in finite-volume quotients of  $\mathrm{PSL}(2, \mathbb{R})$ . *Journal of Modern Dynamics* **10**(2016), 1–21.

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, COLUMBUS, OH 43210  
*E-mail address:* `khalil.37@osu.edu`